STATISTICAL RESEARCH REPORT No 2
Institute of Mathematics 1972
University of Oslo

ON TESTING OF INDEPENDENCE IN $2 \times r$ TABLES BY STEPWISE TEST-PROCEDURES

by

Jens Olav Tveter

Contents

0. Introduction

A statistician sometimes has to deal with problems involving two factors, each having a given number of levels. He wants to test whether the two factors are independent. The usual test-procedure is the so-called test of independence (see [6], chap.XV. 2.3 or other text-books in statistics). This is, however, no exact test, and hence there are certain situations, mostly where the number of observations is small, where the test is less good. In any case, it will be of interest to try to find an exact test.

In the case where one of the factors has only 2 levels, we shall find an exact test for the hypothesis of independence. In that case we arrange the observations in a $2 \times r$ table (see (3.1)).

We will show, that the problem of independence of the factors in a $2 \times r$ table can be put on a regular Darmois-Koopman form, by suitable definitions and change of variables. The hypothesis of independence will then be equivalent to $r - 1$ parameters $\varkappa_1, \dots, \varkappa_{r-1}$ being zero. It is, however, difficult to find an explicit, simultaneous test for this hypothesis. Therefore, we will apply a stepwise testprocedure on the problem, testing one parameter at the time. Thus we may find power-optimal test-procedures on each step, and these procedures will be explicit and relatively easy to find (by means of a computer).

We shall apply two different methods to the problem. The first one is due to T.W. Anderson [3] (also see [1], ch. 1.4.C), and will be referred to as the "method of reduced parameter-space", because parameters not found significantly different from zero, will be considered being exactly zero on the following steps, thus reducing the dimension of the parameter-space by one for each step. This method is relatively simple to apply to most problems, and it often gives simple test-criterions on each step.

The assumption to reduce the dimension of the parameter-space when a parameter is not found $\neq 0$, is difficult, even impossible to introduce in many situations. It is,however, not considered here, we only work on the method in our model.

The other method is due to E.L. Lehman [4], and we will call it the "method of retained parameter space". In this case, we do not consider an accepted hypothesis as correct; on the following steps we will consider the condtional observation-space, given that the hypothesis on the preceding steps have been accepted.

Both methods described above have originally been designed to work with so-called multiple decision problems;we will here, however, only be interested in two decisions: rejection of the hypothesis of independence or saying nothing.

The method of the reduced parameter-space gives an exact test for a given level ϵ. The power-optimality on each step will not, however, lead to power-optimality of the composite test-procedure. But, the latter will be optimal in another way, it will have the maximum performance-function among all performance-unbiased procedures (see [1], pp. 31-35).

The second method is also an exact method, but I have not been able to determine the exact level of the test-procedure, only an upper limit. This is, of course, a serious limitation for practical applications.

Both methods have been programmed and run on a computer, and numerical examples will be found in the appendix.

1. Some theory of regular Darmois-Koopman classes of distributions.

We assume a regular Darmois-Koopman (short D-K) class of distributions, that is, a class where the density (or point-probability) may be written as:

$$dP = A(\tau)e^{\sum_{j=1}^{s}\tau_j Y_j(x)}\, dP_o \tag{1.1}$$

where $\tau_1,\ldots,\tau_s$ are linearly independent and assume their values in an open set Ω , and where $Y_1,\ldots,Y_s$ are linearly independent.

We then infer $m \leq s$ new parameters by

$$\sigma_i = \sum_{j=1}^{s} a_{ij}\tau_j \quad ; \quad i = 1,2,\ldots,m \tag{1.2}$$

Then the following question occurs: When will the class of distributions including the new parameters $\sigma_1,\ldots,\sigma_m$ be regular? The following theorem is easily proved:

Theorem 1: Let $\tau_1,\ldots,\tau_s$ be parameters in a regular D-K class of distributions given by (1.1). Let $\sigma_1,\ldots,\sigma_m$ be $m \leq s$ new parameters given by the linear transformations of $\tau_1,\ldots,\tau_s$ as in (1.2), where the a_{ij} ; $i = 1,2,\ldots,m$, $j = 1,2,\ldots,s$ are such that the vectors

$$\begin{pmatrix} a_{11} \\ \vdots \\ a_{1s} \end{pmatrix}, \begin{pmatrix} a_{21} \\ \vdots \\ a_{2s} \end{pmatrix}, \ldots \begin{pmatrix} a_{m1} \\ \vdots \\ a_{ms} \end{pmatrix}$$

are linearly independent.

Then $\sigma_1,\ldots,\sigma_m$ will be parameters in a regular D-K class of distributions.

From Theorem 1 follows:

Corollary 1: Same assumptions as in Theorem 1, but now let $\sigma_i = \sum_{j=1}^{s} a_{ij}\tau_j + b_i = \sigma_i' + b_i$; $i = 1,2,\ldots,m$. Then the class of distributions with the new parameters is also regular.

Example 1: As an example, consider the multinomial distribution:

$$f(x_1,\ldots,x_n) = \Pr\left(\bigcap_{i=1}^{n} (X_i = x_i)\right) = k_n(x)p_1^{x_1}\cdot\ldots\cdot p_r^{x_r} \tag{1.3}$$

where $\sum_{i=1}^{r} x_i = n$, $\sum_{i=1}^{r} p_i = 1$, $k_n(x) = \dfrac{n!}{x_1!x_2!\ \ldots\ x_r!}$

To put the distribution on a D-K form, we let the distribution P_o correspond to $p_i = \frac{1}{r}$; $i = 1,2,\ldots,r$, that is

$dP_o = k_n(x)(\frac{1}{r})^n d\mu$, μ being the counting-measure.

Then we get:

$$dP = r^n e^{\sum_{j=1}^{r} x_j \ln p_j} dP_o \tag{1.4}$$

This expression for dP is not a regular one, since $\sum_{j=1}^{r} x_j = n$, that is $x_1,\ldots,x_r$ are linearly dependent. We therefore express $x_r = n - x_1 - \ldots - x_{r-1}$, and putting this into (1.4) we get:

$$dP = r^n e^{\sum_{j=1}^{r-1} x_j \ln p_j + (n-x_1-\ldots-x_{r-1})\ln p_r} dP_o$$

$$= (r\cdot p_r)^n e^{\sum_{j=1}^{r-1} x_j \ln p_j/p_r} dP_o = A(\tau)e^{\sum_{j=1}^{r} \tau_j x_j} dP_o \tag{1.5}$$

where $\tau_j = \ln \dfrac{p_j}{1-p_1-\ldots-p_{r-1}}$; $j = 1,2,\ldots,s$; $s = r-1$

$$A(\tau) = [r(1-p_1-\ldots-p_{r-1})]^n$$

It is now obvious that $p_1,\ldots,p_{r-1}$ vary in an open set, therefore $\tau_1,\ldots,\tau_s$ will also vary in an open set, since the

τ_j's are continuous functions of $p_1,\dots,p_{r-1}$; since both $\tau_1,\dots,\tau_s$ and $x_1,\dots,x_s$ $(s=r-1)$ obviously are linearly independent, the class (1.5) is regular.

Theorem 1 now says that all linearly independent linear transformations of the type:

$$\sigma_i = \sum_{j=1}^{s} a_{ij}\tau_j = \sum_{j=1}^{s} a_{ij}\ln\frac{p_j}{p_r} = \sum_{j=1}^{s} a_{ij}\ln p_j - \sum_{j=1}^{s} a_{ij}\ln p_r =$$

$$= \sum_{j=1}^{s} a_{ij}\ln p_j + a_{ir}\ln p_r = \sum_{j=1}^{r} a_{ij}\ln p_j \quad ; \quad i = 1,2,\dots,m$$

where $a_{ir} = -\sum_{j=1}^{s} a_{ij}$; $i = 1,2,\dots,m$; $1 \le m \le s$

will be parameters of a regular D-K class of distributions.

We state this result as a theorem, because of the usefullness in the following chapter.

<u>Theorem 2</u>: Assume a multinomial class of distributions given by (1.3). Define $m \le s = r-1$ new parameters by:

$$\sigma_i = \sum_{j=1}^{r} a_{ij}\ln p_j \quad ; \quad i = 1,2,\dots,m$$

where:

i) $\sum_{j=1}^{r} a_{ij} = 0$; $i = 1,2,\dots,m$

ii) $\{a_{11},\dots,a_{1s}\}'$, $\{a_{21},\dots,a_{2s}\}'$, ..., $\{a_{m1},\dots,a_{ms}\}'$

are m linearly independent vectors.

Then $\sigma_1,\dots,\sigma_m$ are regular D-K parameters.

<u>Example 2</u>: Consider the following transformation of the parameters in a regular D-K class of distributions given by (1.1):

$$\begin{aligned}
\sigma_1 &= \tau_1 \\
\sigma_i &= \tau_i - \tau_{i-1} \quad ; \quad i = 2,3,\ldots,m \\
\sigma_{m+1} &= \tau_{m+1} \\
\sigma_{m+2} &= \tau_{m+2} \\
\sigma_j &= \tau_j - \tau_{j-1} \quad ; \quad j = m+3,m+4,\ldots,s
\end{aligned}$$

where m is an integer between 2 and $s-3$.

The matrix associated with the transformation is:

$$\begin{array}{|cccccc|c|cccccc|l}
1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 \\
-1 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 2 \\
 & \vdots & & & & & & & & & & & \vdots \\
0 & 0 & 0 & \cdots & -1 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & m \\
\hline
0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 & 0 & \cdots & 0 & 0 & m+1 \\
\hline
0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 & 0 & \cdots & 0 & 0 & m+2 \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & -1 & 1 & \cdots & 0 & 0 & m+ \\
 & \vdots & & & & & & & & & & & \vdots \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & -1 & 1 & s
\end{array}$$

It is seen that this matrix has full rank, hence the rows are linearly independent, and according to theorem 1 the class with the new parameters is regular (in the theorem the last column does not count, but the rows are still linearly independent).

2. Testing of independence in a $2 \times r$ table by stepwise testprocedures.

2.1 Definition of independence

In the ordinary 2×2 table to the right we define independence of the two factors A and B by the expression

B \ A	A	A*
B	p_1	p_2
B*	q_1	q_2

$$P(A \cap B) = P(A) \cdot P(B) \text{ , that is } p_1 = (p_1+p_2)(p_1+q_1)$$

or $p_1 q_2 = p_2 q_1$

Now consider the $2 \times r$ table given by (2.1). The symbols in the different boxes give the probability that an observation falls in that box.

B \ A	A_1	A_2		A_{r-1}	A_r	
B	p_1	p_2		p_{r-1}	p_1	p
B^*	q_1	q_2		q_{r-1}	q_1	q

(2.1)

Here the factor A assumes r different levels, $A_1, \ldots, A_r$, while B assumes only two, B and B* (that is, not B).

We will now define independence of A and B in (2.1) in the following way: The probability that a given combination of A and B levels occurs, is equal to the product of the marginal probabilities that the two levels occur, or explicitely:

$(p_1+p_2+\ldots+p_r)(p_i+q_i) = p_i$ i ' $1,2,\ldots,r$

(2,2)

$(q_1+q_2+\ldots+q_r)(p_i+q_i) = q_i$; $i = 1,2,\ldots,r$

The number of condition may be reduced. We immediately see that

$$(p_1+\dots+p_r)(p_i+q_i) = p_i \qquad i = 1,2,\dots,r$$

and

$$p_1+\dots+p_r + q_1+\dots+q_r = 1$$

lead to $(q_1+\dots+q_r)(p_i+q_i) = q_i$; $i = 1,2,\dots,r$, that is, the second line in (2.2) may be deduced from the first.

Further, we see that

$$\sum_{i=1}^{r-1}(p_1+\dots+p_r)(p_i+q_i) = (p_1+\dots+p_r)\sum_{i=1}^{r-1}(p_i+q_i) =$$

$$= (p_1+\dots+p_r)(1-p_r-q_r) = (p_1+\dots+p_r) - (p_1+\dots+p_r)(p_r+q_r)$$

or

$$\sum_{i=1}^{r-1}[(p_1+\dots+p_r)(p_i+q_i) - p_i] = p_r - (p_1+\dots+p_r)(p_r+q_r)$$

that is $(p_1+\dots+p_r)(p_i+q_i) = p_i$; $i = 1,2,\dots,r-1 \implies$

$$\implies (p_1+\dots+p_r)(p_r+q_r) = p_r \,.$$

We have thus shown that

$$(p_1+\dots+p_r)(p_i+q_i) = p_i \; ; \quad i = 1,2,\dots,r-1 \tag{2.3}$$

is a necessary and sufficient condition for (2.2). To obtain an analoguous definition with the 2×2 table, we shall hence define independence of the factors in a 2×2 table by (2.3).

Now we will show that:

$$\begin{array}{cccccccc}
p_1q_2 = p_2q_1 & p_1q_3 = p_3q_1 & \cdot & \cdot & \cdot & \cdot & \cdot & p_1q_r = p_rq_1 \\
\ddots & p_2q_3 = p_3q_2 & \cdot & \cdot & \cdot & \cdot & \cdot & p_2q_r = p_rq_2 \\
 & & \ddots & & & & & \vdots \\
 & & & & & & \ddots & p_{r-1}q_r = p_rq_{r-1}
\end{array} \tag{2.4}$$

is necessary and sufficient for independence.

Assume that (2.4) is true:

Then

$$(p_1+\ldots+p_r)(p_i+q_i) = (p_1+\ldots+p_r)p_i + p_1q_i + p_2q_i+\ldots+p_rq_i =$$
$$= (p_1+\ldots+p_r)p_i + p_iq_1 + p_iq_2+\ldots+p_iq_r =$$
$$= (p_1+\ldots+p_r+q_1+\ldots+q_r)p_i = p_i \;,$$ that is, (2.3), since i was arbitrary.

Now assume that A and B are independent, that is (2.3). Then look at

$$\frac{p_iq_j}{p_jq_i} = \frac{(p_1+\ldots+p_r)(p_i+q_i)(q_1+\ldots+q_r)(p_j=q_j)}{(p_1+\ldots+p_r)(p_j+q_j)(q_1+\ldots+q_r)(p_i+q_i)} = 1$$

which gives (3.4), since i,j were arbitrary.

We see, however, that (2.4) has $\frac{r(r-1)}{2}$ conditions on the parameters, while the definition (2.3) has only $r-1$. Hence it is possible to chose a "base" consisting of $r-1$ functionally independent conditions from (2.4). These may be picked out in several ways.

One of the sets is the diagonal in the matrix (2.4), which we will take a closer look at.

We have

$$p_1q_2 = p_2q_1 \quad p_2q_3 = p_3q_2 \quad \cdots \qquad p_{r-1}q_r = p_rq_{r-1} \tag{2.5}$$

These $r-1$ expressions are functionally independent (that is,none of them may be found as functions of the others), while for arbitrary $i < j-1$ we have:

$$p_iq_j = \frac{p_{i+1}q_i}{q_{i+1}}\cdot q_j = \frac{p_{i+2}\cdot q_{i+1}}{q_{i+2}}\cdot\frac{q_i}{q_{i+1}}\cdot q_j = \frac{p_{i+2}\cdot q_i}{q_{i+2}}\cdot q_j =$$

$$= \cdots = \frac{p_j}{q_j}\cdot q_i\cdot q_j = p_jq_i \;;$$ that is the complete matrix (2.4) may be generated from (2.5), by repeatedly using the fact that

$$p_i = \frac{p_{i+1} \cdot q_i}{q_{i+1}} \quad .$$

The conditions (2.5) have an interesting practical meaning. $p_1 q_2 = p_2 q_1$ is just independence of (A_1, A_2) and B, $p_2 q_3 = p_3 q_2$ means that (A_2, A_3) and B are independent etc., that is we have "reduced" the independence in the $2 \times r$ table to independence in $r-1$ 2×2 tables.

Finally, we will rewrite (2.5) in a way that suits the following. Define

$$\rho_i = \ln \frac{p_i q_{i+1}}{p_{i+1} q_i} \quad ; \qquad i = 1,2,\dots,r-1 \tag{2.6}$$

Then (2.5) is equivalent to

$$\rho_1 = \rho_2 \quad \dots \quad = \rho_{r-1} = 0 \tag{2.7}$$

Notice that the ρ_i's defined by (2.6) are linear expressions of the type

$$\rho_i = \sum_{j=1}^{r} a_{ij} \ln p_j + \sum_{j=1}^{r} b_{ij} \ln q_j \quad ; \quad i = 1,2,\dots,r-1$$

where $\sum_{j=1}^{r} (a_{ij} + b_{ij}) = 0$; $\qquad i = 1,2,\dots,r-1$

But then we have the situation of theorem 2 in chapter 1, where we see that condition i) is fulfilled. All we have to show is that the vectors

$$c_i' = \{a_{i1}, \dots, a_{ir}, b_{i1}, \dots, b_{ir-1}\}' \quad ; \quad i = 1,2,\dots,r-1$$

are linearly independent. Then the parameters $\rho_1, \dots, \rho_{r-1}$ are regular D-K parameters.

To show this, consider the matrix of rows c_i, $i = 1,2,\dots,r-1$

$$\left(\begin{array}{cccccccccc|cccccccc}
1 & -1 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & & 1 & -1 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\
0 & 1 & -1 & 0 & \dots & 0 & 0 & 0 & 0 & & 0 & 1 & -1 & 0 & \dots & 0 & 0 & 0 & 0 \\
\vdots & & & & & & & & & & & & & & & & & & \\
0 & 0 & 0 & 0 & \dots & 0 & 1 & -1 & 0 & & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & -1 \\
0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & -1 & & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 1
\end{array}\right)$$

That this matrix has full rank is easily seen (by generally adding the $j-1$ preceeding rows to the j row, $j = 2,3,\dots,r$, and keep the first row unchanged).

We have thus shown: $\rho_1,\dots,\rho_{r-1}$ are regular D-K parameters, and the problem of independence in a $2 \times r$ table may be put on a regular D-K form.

2.2 Transformation of the problem to D-K form

We observe tha factors A and B during N trials, and arrange the results in the table below:

B \ A	A_1	A_2	A_3	– – – – – –	A_{r-1}	A_r	
B	x_1	x_2	x_3	– – – – – –	x_{r-1}	x_r	N-w
B^*	y_1	y_2	y_3	– – – – – –	y_{r-1}	y_r	w
	u_1	u_2	u_3		u_{r-1}	u_r	N

(2.8)

According to the usual assumptions for trials of this type, $(X_1,\dots,X_r,Y_1,\dots,Y_r)$ have a multinomial distribution with probabilities given by (2.1), that is

$$\Pr\left(\bigcap_{i=1}^{r}(X_i = x_i \cap Y_i = y_i)\right) = f = k_N(x,y)p_1^{x_1}\dots p_r^{x_r}q_1^{y_1}\dots q_r^{y_r} ,$$

$$\text{for} \quad \sum_{i=1}^{r}(x_i+y_i) = N , \quad 0 < p_i < 1 , \quad 0 < q_i < 1$$

where $k_N(x,y) = \dfrac{N!}{x_1!\ldots x_r!y_1!\ldots y_r!}$

In chapter 2.1 we have shown that the problem of independence in a $2 \times r$ table may be put on a regular D-K form. To reach this form, it is necessary to make a small detour.

We start with the following $r-1$ conditions

$$p_1q_2 = p_2q_1 \ , \ p_1q_3 = p_3q_1 \ , \ \ldots \ , \ p_1q_r = p_rq_1 \qquad (2.9)$$

These conditions are obviously functionally independent, and (2.9) is equivalent to (2.4). We shall prove the latter.

Assume (2.9) to be true: Let $1 < i < j$, and look at p_iq_j . From (2.9) it follows that

$$p_i = \frac{p_1q_i}{q_1} \ , \quad q_j = \frac{p_j \cdot q_1}{p_1} \ , \quad \text{hence}$$

$$p_iq_j = \frac{p_1q_i}{q_1} \cdot \frac{p_jq_1}{p_1} = p_jq_i \ , \text{ which proves (2.4)}$$

We have previously shown that the opposite is true.

As in chapter 2.1, we infer logarithmic parameters. Let

$$\eta_1 = \ln\frac{p_1q_2}{p_2q_1} \ , \ \eta_2 = \ln\frac{p_1q_3}{p_3q_1} \ \ldots \ \eta_{r-1} = \ln\frac{p_1q_r}{p_rq_1} \qquad (2.10)$$

Then $\eta_1,\ldots,\eta_{r-1}$ are regular D-K parameter from similar reasons as in 2.1.

The probability function f was given by:

$$f = k_N(x,y)p_1^{x_1}\cdot\ldots\cdot p_r^{x_r}q_1^{y_1}\cdot\ldots\cdot q_r^{y_r} =$$

$$= k_N(x,y)\left(\frac{p_1q_2}{p_2q_1}\right)^{y_2}\cdot\ldots\cdot\left(\frac{p_1q_r}{p_rq_1}\right)^{y_r}p_1^{x_1}\cdot\ldots\cdot p_r^{x_r}q_1^{y_1}p_1^{-y_2-\ldots-y_r} \cdot$$

$$\cdot\ q_1^{y_2+\ldots+y_r}\cdot p_2^{y_2}\cdot\ldots\cdot p_r^{y_r} =$$

$$= k_N(x,y)\, e^{\sum_{j=1}^{r-1} y_{j+1}\cdot \eta_j} \cdot p_1^{u_1-w} \cdot p_2^{u_2} \cdot \ldots \cdot p_r^{u_r} \cdot q_1^{w} =$$

$$= k_N(x,y)p_1^N\, e^{\sum_{j=1}^{r-1} y_{j+1}\eta_j} \left(\frac{q_1}{p_1}\right)^{w} \left(\frac{p_2}{p_1}\right)^{u_2} \cdot \ldots \cdot \left(\frac{p_r}{p_1}\right)^{u_r} =$$

$$= k_N(x,y)p_1^N\, e^{\sum_{j=1}^{r-1} y_{j+1}\eta_j + w\cdot\rho_1 + \sum_{j=2}^{r} u_j\rho_j} \qquad (2.11)$$

where $\rho_1 = \ln \frac{q_1}{p_1}$, $\rho_j = \ln \frac{p_j}{p_1}$; $i = 2,3,\ldots,r$,

and $u_j = x_j + y_j$; $j = 1,2,\ldots,r$, $w = y_1+\ldots+y_r$.

Now consider the coefficient-matrix of the transformation from $(\ln p_1, \ldots, \ln q_r)$ to $(\eta_1,\ldots,\eta_{r-1},\rho_1,\ldots,\rho_r)$, which is shown below. The $2r-1$ rows are seen to be linearly independent. Each row has the same number of $+1$ and -1 , hence Theorem 2 in chapter 1 may be used to show that $\eta_1,\ldots,\eta_{r-1},\rho_1,\ldots,\rho_r$ are parameters of a regular D-K class (the theorem requires that the rows are independent when e.g. the last column is removed, which is trivial).

The coefficient-matrix is as follows

$$\begin{array}{c}
\left[\begin{array}{cccccc|cccccc}
1 & -1 & 0 & \cdots & & 0 & -1 & 1 & 0 & \cdots & & 0 \\
1 & 0 & -1 & \cdots & & 0 & -1 & 0 & 1 & \cdots & & 0 \\
\vdots & & & \ddots & & & & & & \ddots & & \vdots \\
1 & 0 & 0 & \cdots & & -1 & -1 & 0 & 0 & \cdots & & 1 \\
\hline
-1 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 & & \cdots & & 0 \\
\hline
-1 & 1 & 0 & \cdots & & 0 & 0 & & & \cdots & & 0 \\
-1 & 0 & 1 & \cdots & & & \vdots & & & & & \vdots \\
\vdots & & & \ddots & & & & & & & & \\
-1 & 0 & 0 & \cdots & & 1 & 0 & & & \cdots & & 0
\end{array}\right]
\begin{array}{l}
\left.\begin{array}{c} \\ \\ \\ \\ \end{array}\right\} r-1 \\
\\
\left.\begin{array}{c} \\ \\ \\ \\ \end{array}\right\} r-1
\end{array}
\\
\underbrace{\hspace{6cm}}_{r}\underbrace{\hspace{6cm}}_{r}
\end{array}$$

Now we infer new variables $C_1,\dots,C_{r-1}$ for $Y_2,\dots,Y_r$ by:

$$C_i = \sum_{j=i+1}^{r} Y_j \quad ; \quad i = 1,2,\dots,r-1 \tag{2.12}$$

This gives:

$$Y_i = C_{i-1} - C_i \quad ; \quad i = 2,\dots,r-1$$
$$Y_r = C_{r-1}$$

Similarly we transform the $U_1,\dots,U_{r-1}$ to $V_1,\dots,V_{r-1}$ by

$$V_i = \sum_{j=i+1}^{r} U_j \quad ; \quad i = 1,2,\dots,r-1 \tag{2.13}$$

which is equivalent to:

$$U_i = V_{i-1} - V_i \quad ; \quad i = 2,3,\dots,r-1$$
$$U_r = V_{r-1}$$

If we let $C_r = V_r = 0$, we may write $Y_r = C_{r-1} - C_r$, $U_r = V_{r-1} - V_r$, and hence:

$$\sum_{j=1}^{r-1} Y_{j+1} \cdot \eta_j + \sum_{j=2}^{r} U_j \rho_j =$$

$$= \sum_{j=1}^{r-1} \eta_j (C_j - C_{j+1}) + \sum_{j=2}^{r} \rho_j (V_{j-1} - V_j) =$$

$$= \sum_{j=1}^{r-1} \eta_j C_j - \sum_{j=2}^{r} \eta_{j-1} C_j + \sum_{j=1}^{r-1} \rho_{j+1} V_j - \sum_{j=2}^{r} \rho_j V_j =$$

$$= \eta_1 C_1 + \sum_{j=2}^{r-1} (\eta_j - \eta_{j-1}) C_j - \eta_{r-1} C_r + \rho_2 V_1 + \sum_{j=2}^{r-1} (\rho_{j+1} - \rho_j) V_j - \rho_r V_r =$$

$$= \eta_1 C_1 + \sum_{j=2}^{r-1} (\eta_j - \eta_{j-1}) C_j + \rho_2 V_1 + \sum_{j=2}^{r-1} (\rho_{j+1} - \rho_j) V_j ,$$

since $V_r = C_r = 0$.

Now put $\varkappa_1 = \eta_1 = \ln \dfrac{p_1 q_2}{p_2 q_1}$,

$$\varkappa_i = \eta_i - \eta_{i-1} = \ln \frac{\dfrac{p_1 q_{i+1}}{p_{i+1} q_1}}{\dfrac{p_1 q_i}{p_i q_1}} = \ln \frac{p_i q_{i+1}}{p_{i+1} q_i} \; ; \quad i = 2,3,\dots,r-1$$

Further let $\tau_o = \rho_1 = \ln \dfrac{q_1}{p_1}$, $\tau_1 = \rho_2 = \ln \dfrac{p_2}{p_1}$,

$$\tau_i = \rho_{i+1} - \rho_i = \ln \frac{\dfrac{p_{i+1}}{p_1}}{\dfrac{p_i}{p_1}} = \ln \frac{p_{i+1}}{p_i} \; ; \qquad i = 2,3,\dots,r-1$$

Then (2.11) becomes:

$$dP = (2rp_1)^N \, e^{\sum\limits_{j=1}^{r=1} \varkappa_i c_i + \tau_o W + \sum\limits_{j=1}^{r-1} \tau_j v_j} dP_o \tag{2.14}$$

or, by letting $V_o = W = \sum\limits_{j=1}^{r} U_j$

$$dP = (2rp_1)^N \, e^{\sum\limits_{j=1}^{r-1} \varkappa_i c_i + \sum\limits_{j=o}^{r-1} \tau_j v_j} dP_o \tag{2.14a}$$

Here P_o corresponds to $p_i = q_j = \frac{1}{2r}$, $\forall i,j$, that is

$$dP_o = k_N(x,y)\left(\frac{1}{2r}\right)^N d\mu \tag{2.15}$$

where μ is the counting-measure.

We see that the class (2.14) is regular, from example 2, chapter 1, considering the transformation of the parameters.

2.3.1 The test-procedure following from the principle of reduced parameter space.

The hypothesis of independence of A and B is equivalent to

$$H:\ \varkappa_1 = \varkappa_2 = \dots = \varkappa_{r-1} = 0$$

We will test this hypothesis stepwise. The principle of reduced parameter space says the following:

First we test H_1: $\varkappa_{r-1} = 0$, if H_1 is accepted, then test H_2: $\varkappa_{r-2} = 0$ under the assumption of $\varkappa_{r-1} = 0$, if H_2 is rejected, stop the procedure. If H_2 is accepted, then test H_3: $\varkappa_{r-3} = 0$, putting $\varkappa_1 = \varkappa_2 = 0$, otherwise stop ; etc.

The process will stop either if one of the hypothesis H_i: $\varkappa_{r-i} = 0$ (assuming $\varkappa_{r-1} = \dots = \varkappa_{r-i+1} = 0$) is rejected, then the hypothesis of independence, H is rejected; or if all H_i's are accepted, in that case the hypothesis H is accepted.

Step 1.

We will test H_1: $\varkappa_{r-1} = 0$ in the distribution (2.14). Since (2.14) describes a regular D-K class of distributions, a well-known theorem (see e.g. [7] chapter 4.4) states that the uniformly most powerful unbiased test (UMPU) for H_1 against a two-sided alternative is given by

$$\delta_1(C_{r-1},T_1) = \begin{cases} 1 & \text{if } C_{r-1} < k_1(T_1) \text{ or } C_{r-1} > h_1(T_1) \\ \gamma_{11}(T_1) & \text{if } C_{r-1} = k_1(T_1) \\ \gamma_{21}(T_1) & \text{if } C_{r-1} = h_1(T_1) \\ 0 & \text{otherwise} \end{cases} \tag{2.16}$$

where $T_1 = (C_1,\dots,C_{r-2},W,V_1,\dots,V_{r-1})$.

The constants k_1, h_1, γ_{11}, γ_{21} (depending on T_1) are determined as to fulfill the following conditions

$$E_o(\delta_1 | T_1) = \epsilon_1$$

$$E_o(\delta_1 \cdot C_{r-1} | T_1) = \epsilon_1 \cdot E_o(C_{r-1} | T_1)$$

where $E_o(\cdot | T_1)$ is the conditional expectation given T_1 under P_o .

We will hence have to find the conditional probability function of $C_{r-1} | T_1$ under P_o , that is, when all parameters in (2.14) are equal to zero.

We define

$$g_1(c_{r-1};t_1) = Pr(C_{r-1} = c_{r-1} | T_1 = t_1) = \frac{Pr(C_{r-1}=c_{r-1} \cap T_1=t_1)}{Pr(T_1=t_1)} =$$

$$= \frac{f(c_{r-1}, t_1)}{f_1(t_1)} \tag{2.17}$$

Under P_o we have

$$Pr(\bigcap_{i=1}^{r} (X_i = x_i \cap Y_i = y_i)) = \frac{N!}{x_1! \dots x_r! y_1! \dots y_r!} (\frac{1}{2r})^N$$

Transforming to the new variables $(C_1, \dots, C_{r-1}, W, V_1, \dots, V_{r-1}) = (C_{r-1}, T_1)$ by (2.12) and (2.13), gives, since the transformations are one-one

$$f(c_{r-1}, t_1) = \frac{N!}{(N-v_1(w_1-(1))!(v_1-v_2-(c_1-(2))! \dots (v_{r-1}c_{r-1})!} \cdot \\ \cdot \frac{1}{(w-c_1)!(c_1-c_2)! \dots c_{r-1}!} \cdot (\frac{1}{2r})^N \tag{2.18}$$

We further need

$$f_1(t_1) = f_1(c_1, \dots, c_{r-2}, w, v_1, \dots, v_{r-1}) = \sum_{c_{r-1}=0}^{c_{r-2}} f(c_{r-1}, t_1)$$

To calculate the sum on the right hand side, first consider the sum of the terms in (2.18) which contains c_{r-1}

$$\sum_{c_{r-1}=0}^{c_{r-2}} [(v_{r-2}-v_{r-1}-(c_{r-2}-c_{r-1}))!(v_{r-1}-c_{r-1})!(c_{r-2}-c_{r-1})!c_{r-1}!]^{-1} =$$

$$= \frac{1}{(v_{r-2}-v_{r-1})!v_{r-1}!} \sum_{c_{r-1}=0}^{c_{r-2}} \binom{v_{r-2}-v_{r-1}}{c_{r-2}-c_{r-1}}\binom{v_{r-1}}{c_{r-1}} = \binom{v_{r-2}}{c_{r-2}} \frac{1}{(v_{r-2}-v_{r-1})!v_{r-1}!}$$

using a result of the hypergeometric distribution.

This gives

$$f_1(t_1) = \frac{N!}{(N-v_1-(w-c_1))!(v_1-v_2-(c_1-c_2))!\dots(v_{r-3}-v_{r-2}-(c_{r-3}-c_{r-2}))!}$$

$$\frac{1}{(w-c_1)!(c_1-c_2)!\dots(c_{r-3}-c_{r-2})!v_{r-1}!(v_{r-2}-v_{r-1})!}\binom{v_{r-2}}{c_{r-2}}\frac{1}{(2r)^N} \tag{2.19}$$

and hence, dividing (2.18) by (2.19), and arranging

$$g_1(c_{r-1};t_1) = \frac{v_{r-1}!(v_{r-2}-v_{r-1})!}{(v_{r-2}-v_{r-1}-(c_{r-2}-c_{r-1}))!(v_{r-1}-c_{r-1})!(c_{r-2}-c_{r-1})!c_{r-1}!}\binom{v_{r-2}}{c_{r-2}}^{-1} =$$

$$= \frac{v_{r-1}!}{(v_{r-1}-c_{r-1})!c_{r-1}!} \cdot \frac{(v_{r-2}-v_{r-1})!}{(v_{r-2}-v_{r-1}-(c_{r-2}-c_{r-1}))!} \cdot \binom{v_{r-2}}{c_{r-2}}^{-1} =$$

$$= \binom{v_{r-1}}{c_{r-1}}\binom{v_{r-2}-v_{r-1}}{c_{r-2}-c_{r-1}} \cdot \binom{v_{r-2}}{c_{r-2}}^{-1}; \tag{2.19a}$$

if $\max(c_{r-2}+v_{r-1}-v_{r-2},0) \leq c_{r-1} \leq \min(v_{r-1},c_{r-2})$

$g_1(c_{r-1};t_1) = 0$ otherwise

This result may be understood as to test independence of (A_{r-1},A_r) and B given $A_1,A_2,\dots,A_{r-2}$.

Step i

We assume that the proceding hypothesis $H_1,\dots,H_{i-1}$ are all accepted, that is we assume $\varkappa_{r-1}=\dots=\varkappa_{r-i+1} = 0$. Then we want to test $H_i\colon \varkappa_{r-i} = 0$.

The corresponding distribution function put on D-K form is found from (2.14) by putting $\varkappa_{r-1}=\dots=\varkappa_{r-i+1} = 0$ and summing over all $(c_{r-i+1},\dots,c_{r-1})$. Then we find

$$dP^{(i-1)} = (2r\,p_1)^N e^{\sum_{j=1}^{r-i} \varkappa_j c_j + \tau_o W + \sum_{j=o}^{r-1} \tau_j v_j} dP_o^{(i-1)}$$

where $P_o^{(i-1)}$ is the distribution of $(C_1,\dots,C_{r-i},W,V_1,\dots,V_{r-1})$ where $\varkappa_1=\dots=\varkappa_{r-i} = 0 = \tau_o=\dots=\tau_{r-1}$.

The form is again regular, and using the same result as on step 1, we find that the UMPU test for $H_i\colon \varkappa_{r-i} = 0$ against a two-sided alternative is given by

$$\delta_i(C_{r-i},T_i) = \begin{cases} 1 & \text{if } C_{r-i} < k_i(T_i) \text{ or } C_{r-i} > h_i(T_i) \\ \gamma_{1i}(T_i) & \text{if } C_{r-i} = k_i(T_i) \\ \gamma_{2i}(T_i) & \text{if } C_{r-i} = h_i(T_i) \\ 0 & \text{otherwise} \end{cases} \tag{2.20}$$

where $T_i = (C_1,\dots,C_{r-i+1},W,V_1,\dots,V_{r-1})$ and where the constants $k_i, h_i, \gamma_{1i}, \gamma_{2i}$ (functions of T_i) are determined by

$$\begin{aligned} E_o(\delta_i|T_i) &= \epsilon_i \\ E_o(\delta_i\cdot C_{r-i}|T_i) &= \epsilon_i\cdot E_o(C_{r-i},T_i) \end{aligned} \tag{2.21}$$

where $E_o(\cdot|T_i)$ is the conditional expectation given T_i in the distribution $P_o^{(i-1)}$.

Now we need the conditional distribution of $C_{r-i}|T_i$ under $P_o^{(i-1)}$. We have that:

$$g_i(c_{r-i};t_i) = Pr(C_{r-i}=c_{r-i}|T_i=t_i) = \frac{Pr(C_{r-i}=c_{r-i} \cap T_i=t_i)}{Pr(T_i=t_i)} =$$

$$= \frac{f_{i-1}(c_{r-i},t_i)}{f_i(t_i)} = \frac{f_{i-1}(t_{i-1})}{f_i(t_i)} \qquad (2.22)$$

We shall now have to prove the general expression for $f_i(t_i) = f_i(c_{r-i-1},t_{i+1})$, namely:

$$f_i(t_i) = \frac{N!}{(N-v_1-(w-c_1))!(v_1-v_2-(c_1-c_2))!\dots(v_{r-i-2}-v_{r-i-1}-(c_{r-i-2}-c_{r-i-1}))!}$$

$$\cdot \frac{1}{(w-c_1)!(c_1-c_2)!\dots(c_{r-i-2}-c_{r-i-1})!} \cdot$$

$$\cdot \frac{\binom{v_{r-i-1}}{c_{r-i-1}}}{v_{r-1}!(v_{r-2}-v_{r-1})!\dots(v_{r-i-1}-v_{r-i})!}\left(\frac{1}{2r}\right)^N \qquad (2.23)$$

for $i = 1,2,\dots,r-2$

$$f_{r-1}(t_{r-1}) = \frac{\binom{N}{w}}{(N-v_1)!} \cdot \frac{N!}{v_{r-1}!(v_{r-2}-v_{r-1})!\dots(v_1-v_2)!}\left(\frac{1}{2r}\right)^N$$

We have shown (2.23) for $i = 1$, see (2.19). Now assume (2.23) to be true for $i = k-1$, that is

$$f_{k-1}(t_{k-1}) = \frac{N!}{(N-v_1-(w-c_1))!(v_1-v_2-(c_1-c_2))!\dots(v_{r-k-1}-v_{r-k}-(c_{r-k-1}-c_{r-k}))!}$$

$$\cdot \frac{1}{(w-c_1)!(c_1-c_2)!\dots(c_{r-k-1}-c_{r-k})!} \cdot$$

$$\frac{\binom{v_{r-k}}{c_{r-k}}}{v_{r-1}!(v_{r-2}-v_{r-1})!(v_{r-3}-v_{r-2})!\dots(v_{r-k}-v_{r-k+1})!}\left(\frac{1}{2r}\right)^N$$

As before, we find

$$f_k(t_k) = \sum_{c_{r-k}=o}^{c_{r-k-1}} f_{k-1}(t_{k-1}) = \sum_{c_{r-k=o}}^{c_{r-k-1}} f_{k-1}(c_{r-k},t_k)$$

We first take the terms depending on c_{r-k} :

$$\sum_{c_{r-k}=o}^{c_{r-k-1}} \frac{\binom{v_{r-k}}{c_{r-k}}}{(v_{r-k-1}-v_{r-k}-(c_{r-k-1}-c_{r-k}))!(c_{r-k-1}-c_{r-k})!} =$$

$$= \sum_{c_{r-k}=o}^{c_{r-k-1}} \binom{v_{r-k-1}-v_{r-k}}{c_{r-k-1}-c_{r-k}} \binom{v_{r-k}}{c_{r-k}} \frac{1}{(v_{r-k-1}-v_{r-k})!} =$$

$$= \frac{\binom{v_{r-k-1}}{c_{r-k-1}}}{(v_{r-k-1}-v_{r-k})!} \quad ; \quad \text{such that}$$

$$f_k(t_k) = \frac{N!}{(N-v_1-(w-c_1))!(v_1-v_2-(c_1-c_2))!\dots(v_{r-k-2}-v_{r-k-1}-(c_{r-k-2}-c_{r-k-1}))!} \cdot$$

$$\cdot \frac{1}{(w-c_1)!(c_1-c_2)!\dots(c_{r-k-2}-c_{r-k-1})!} \cdot$$

$$\cdot \qquad \cdot \frac{\binom{v_{r-k-1}}{c_{r-k-1}}}{v_{r-1}!(v_{r-2}-v_{r-1})!\dots(v_{r-k-1}-v_{r-k})!} \left(\frac{1}{2r}\right)^N$$

proving (2.23) for $i = 1,2,\dots,r-2$.

For $i = r-1$ the expressions get somewhat simpler. We have

$$f_{r-2}(t_{r-2}) = \frac{N!}{(N-v_1-(w-c_1))!} \cdot \frac{1}{(w-c_1)!} \cdot \frac{\binom{v_1}{c_1}}{v_{r-1}!(v_{r-2}-v_{r-1})!\dots(v_1-v_2)!} \left(\frac{1}{2r}\right)^N$$

summing over c_1 , we get

$$f_{r-1}(t_{r-1}) = \sum_{c_1=o}^{w} f_{r-2}(t_{r-2}) =$$

$$\sum_{c_1=o}^{w} \frac{N!}{(N-v_1-(w-c_1))!} \cdot \frac{1}{(w-r_1)!} \cdot \frac{\binom{v_1}{c_1}}{v_{r-1}!(v_{r-2}-v_{r-1})!\dots(v_1-v_2)!} \left(\frac{1}{2r}\right)^N =$$

$$= \sum_{c_1=o}^{w} \frac{\binom{N-v_1}{w-c_1}\binom{v_1}{c_1}}{(N-v_1)!} \cdot \frac{N!}{v_{r-1}!(v_{r-2}-v_{r-1})!\ldots(v_1-v_2)!} \left(\frac{1}{2r}\right)^N =$$

$$= \frac{\binom{N}{w}}{(N-v_1)!} \cdot \frac{N!}{v_{r-1}!(v_{r-2}-v_{r-1})!\ldots(v_1-v_2)!} \left(\frac{1}{2r}\right)^N$$

Then (2.23) is proved (by induction), and the expression for $g_i(c_{r-i},t_i)$ is formed by dividing $f_{i-1}(t_{i-1})$ by $f_i(t_i)$ (by (2.22)). We find; after some work:

$$g_i(c_{r-i};t_i) = \begin{cases} \dfrac{\binom{v_{r-i-1}-v_{r-i}}{c_{r-i-1}-c_{r-i}}\binom{v_{r-i}}{c_{r-i}}}{\binom{v_{r-i-1}}{c_{r-i-1}}} & \text{for } \max(c_{r-i-1}+v_{r-i}-v_{r-i-1},0) \\ & \leq c_{r-i} \leq \min(v_{r-i},c_{r-i-1}) \\ 0 & \text{otherwise} \end{cases} \tag{2.24}$$

being valid for $i = 1,2,\ldots,r-1$. (Remember that $C_o = W$, $V_o = N$.)

We have now given the method in principle, because now we are able to calculate all the constants k_i, h_i, γ_{1i}, γ_{2i} in (2.20), thus being able to carry out the tests. The practical form of the method is given in the next section.

As we have seen, one method reduces the problem of independence in a $2 \times r$ - table to $r-1$ tests of independence in 2×2 tables.

One of the immediate weaknesses of this testprocedure, is that it takes acception of a sub-hypothesis to mean independence in the corresponding 2×2 table, and "adds" all the independent 2×2 - tables to one accumulated table. If we let $A_i \cup A_j$ mean that the levels A_i and A_j are put together, that is, their frequenties are added, the i'th step of the method will test independence in the 2×2 table $(A_{r-i}, A_{r-i+1} \cup \ldots \cup A_r) \times (B, B^*)$, given $A_1, \ldots, A_{r-i+1}$.

The method, however, has one optimal property. More generally, our problem may be formulated as a multiple decision problem, that is, we want to chose between one and only one of the following decisions:

$$d_j : \varkappa \in \Omega_j \quad ; \quad j = 1,2,\ldots,r$$

where $\varkappa = (\varkappa_1, \ldots, \varkappa_{r-1})$, and the Ω_i are defined as follows:

$$\Omega_1 : \varkappa_{r-1} \neq 0$$
$$\Omega_2 : \varkappa_{r-1} = 0 , \quad \varkappa_{r-2} \neq 0$$
$$\vdots$$
$$\Omega_j : \varkappa_{r-1} = \ldots = \varkappa_{r-j+1} = 0 , \quad \varkappa_{r-j} \neq 0$$
$$\vdots$$
$$\Omega_r : \varkappa_{r-1} = \ldots = \varkappa_1 = 0$$

A decision-procedure ψ is defined by a r-tuple $\psi = (\psi_1, \ldots, \psi_r)$, where the ψ_j is the indicator function of accepting d_j, that is $\sum_{j=1}^{r} \psi_j = 1$. ψ_j is of course a function of the sufficient parameters in our D-K class of distributions. The expected value of ψ, $E\psi = (E\psi_1, \ldots, E\psi_r)$ is called the performance function of the method. Let us put some conditions on this performance function:

(i) $$E(\psi_i \mid \varkappa_{r-1} = \ldots = \varkappa_{r-i} = 0) = p_i \quad ; \; i = 1,2,\ldots,r$$

where $p_1, \ldots, p_r$ are given numbers, $p_1, \ldots, p_{r-1}$ small, such that $p_1 + \ldots + p_r = 1$. That means, the probability of stating $d_i : \varkappa_{r-i} \neq 0$, $\varkappa_{r-1} = \ldots = \varkappa_{r-i+1} = 0$ when in fact $\varkappa_{r-1} = \ldots = \varkappa_{r-i} = 0$ is p_i.

(ii) $\quad E(\psi_i \mid \varkappa_{r-1}=\dots=\varkappa_{r-i+1}=0) \geq p_i \qquad i = 1,2,\dots,r-1$

When condition (ii) is fulfilled, the ψ is said to be performance unbiased.

Now, used on the model outlined here, our method can be shown to maximize the performance among all performance-unbiased decision procedures. See [3] or [1].

We will now consider the total level of the test-method. Under this discussion, we give some special definitions. Let F mean that the hypothesis of independence in the $2 \times r$ table is rejected, and let F_i mean that H_i is rejected, while E_i means that it is accepted. Then the total level ϵ may be written:

$$\epsilon = P_r(F) = Pr(\bigcup_{i=1}^{r-1} F_i) = 1 - Pr(\bigcap_{i=1}^{r-1} E_i) \tag{2.25}$$

if the hypothesis is true.

The multiplication law applied on $Pr(\bigcap_{i=1}^{r-1} E_i)$ gives

$$Pr(\bigcap_{i=1}^{r-1} E_i) = Pr(E_1)Pr(E_2 \mid E_1),\dots,Pr(E_{r-1} \mid E_1\cap\dots\cap E_{r-2})$$

where, "given $E_1\cap\dots\cap E_k$" means that the hypothesis $H_1,\dots,H_k$ is accepted. From our preceeding definitions we have, that if the hypothesis H is true, then

$$Pr((E_{k+1} \mid E_1\cap\dots\cap E_k) \mid T_k) = 1 - Pr((F_{k+1} \mid E_1\cap\dots\cap E_k) \mid T_k) = 1-\epsilon_{k+1},$$

exactly, for all possible values of T_k. But then we may sum away the "given T_k", since ϵ_{k+1} is independent of T_k. This gives:

$$Pr(E_{k+1} \mid E_1\cap\dots\cap E_k) = 1 - Pr(F_{k+1} \mid E_1\cap\dots\cap E_k) = 1 - \epsilon_{k+1}$$

for $k = 0,1,2,\dots,r-2$, hence

$$\Pr(\bigcap_{i=1}^{r-1} E_i) = (1-\varepsilon_1)(1-\varepsilon_2)\ldots\ldots(1-\varepsilon_{r-1})$$

which by (2.25) gives

$$\varepsilon = 1-(1-\varepsilon_1)(1-\varepsilon_2)\ldots\ldots(1-\varepsilon_{r-1}) \tag{2.26}$$

The expression also has another form, with an interesting interpretation, which is formed by multiplying out one term at the time, starting from the right.

We get

$$\varepsilon = \varepsilon_1 + (1-\varepsilon_1)\varepsilon_2 + (1-\varepsilon_1)(1-\varepsilon_2)\varepsilon_3 + \ldots + (1-\varepsilon_1)(1-\varepsilon_2),\ldots,(1-\varepsilon_{r-2})\varepsilon_{r-1} \tag{2.27}$$

Here the form $(1-\varepsilon_1)(1-\varepsilon_2),\ldots,(1-\varepsilon_{i-1})\varepsilon_i$ may be interpreted as the probability of rejecting H_i and accepthing $H_1, H_2, \ldots$ $\ldots, H_{i-1}$, which is equal to the probability of rejecting H_i given the acceptance of $H_1,\ldots,H_{i-1}$ multiplied with the probability of the latter.

Hence we have constructed a level ε test-procedure for the hypothesis H consiting of at most $r-1$ steps.

Since the purpose of our test is to decide between dependence or not, it seems natural to use the smme level on each step, that is $\varepsilon_i = \alpha$; $i = 1,2,\ldots,r-1$.

This gives

$$\varepsilon = \alpha + \alpha(1-\alpha) + \ldots + \alpha(1-\alpha)^{r-2} = 1-(1-\alpha)^{r-1} ,$$

from (2.26). Therefore, if we want total level ε , we chose

$$\alpha = 1 - \sqrt[r-1]{1-\varepsilon} \tag{2.28}$$

2.3.2. The practical appearance of the test-procedure

To use the test in practice, we shall put (2.21) on a more explicit form. The distribution in which the expectation $E_o(\cdot|T_i)$ is taken, is the hypergeometric given by (2.24). Using (2.20), we may write (2.21) as

$$\text{(i)}\quad \sum_{c=o}^{k_i(T_i)-1} \frac{\binom{v_{r-i}}{c}\binom{v_{r-i-1}-v_{r-i}}{c_{r-i-1}-c}}{\binom{v_{r-i-1}}{c_{r-i-1}}} + \gamma_{1i}(T_i)\cdot\frac{\binom{v_{r-i}}{k_i(T_i)}\binom{v_{r-i-1}-v_{r-i}}{c_{r-i-1}-k_i(T_i)}}{\binom{v_{r-i-1}}{c_{r-i-1}}} + \gamma_{2i}(T_i)\cdot\frac{\binom{v_{r-i}}{h_i(T_i)}\binom{v_{r-i-1}-v_{r-i}}{c_{r-i-1}-h_i(T_i)}}{\binom{v_{r-i-1}}{c_{r-i-1}}} + \sum_{c=h_i(T_i)+1}^{c_{r-i-1}} \frac{\binom{v_{r-i}}{c}\binom{v_{r-i-1}-v_{r-i}}{c_{r-i-1}-c}}{\binom{v_{r-i-1}}{c_{r-i-1}}} = \varepsilon_i \qquad (2.29)$$

$$\text{(ii)}\quad \sum_{c=o}^{k_i(T_i)-1} c\cdot\frac{\binom{v_{r-i}}{c}\binom{v_{r-i-1}-v_{r-i}}{c_{r-i-1}-c}}{\binom{v_{r-i-1}}{c_{r-i-1}}} + \gamma_{1i}(T_i)k_i(T_i)\cdot\frac{\binom{v_{r-i}}{k_i(T_i)}\binom{v_{r-i-1}-v_{r-i}}{c_{r-i-1}-c_{r-i}}}{\binom{v_{r-i-1}}{c_{r-i-1}}} + \gamma_{2i}(T_i)\cdot h_i(T_i)\cdot\frac{\binom{v_{r-i}}{h_i(T_i)}\binom{v_{r-i-1}-v_{r-i}}{c_{r-i-1}-h_i(T_i)}}{\binom{v_{r-i-1}}{c_{r-i-1}}} + \sum_{c=h_i(T_i)+1}^{c_{r-i-1}} c\cdot\frac{\binom{v_{r-i}}{c}\binom{v_{r-i-1}-v_{r-i}}{c_{r-i-1}-c}}{\binom{v_{r-i-1}}{c_{r-i-1}}} = \qquad (2.30)$$

$$= \varepsilon_i \sum_{c=o}^{c_{r-i-1}} \frac{\binom{v_{r-i}}{c}\binom{v_{r-i-1}-v_{r-i}}{c_{r-i-1}-c}}{\binom{v_{r-i-1}}{c_{r-i-1}}}$$

These equations are typographically incomfortable, therefore, let me for a moment consider the second equation with simpler variable-names.

Let $a = v_{r-i}$, $x = c$, $N = v_{r-i-1}$, $n = c_{r-i-1}$, $\gamma_1 = \gamma_{1i}(T_i)$, $\gamma_2 = \gamma_{2i}(T_i)$, $k = k_i(T_i)$, $h = h_i(T_i)$, $\epsilon = \epsilon_i$. Then the equation becomes:

$$\sum_{x=0}^{k-1} x \cdot \frac{\binom{a}{x}\binom{N-a}{n-x}}{\binom{N}{n}} + \gamma_1 \cdot k \cdot \frac{\binom{a}{k}\binom{N-a}{n-k}}{\binom{N}{n}} + \gamma_2 \cdot h \cdot \frac{\binom{a}{h}\binom{N-a}{n-h}}{\binom{N}{n}} + \sum_{x=h+1}^{n} x \cdot \frac{\binom{a}{x}\binom{N-a}{n-x}}{\binom{N}{n}} =$$

$$= \epsilon \sum_{x=0}^{n} x \cdot \frac{\binom{a}{x}\binom{N-a}{n-x}}{\binom{N}{n}}$$

Instead of $x \cdot \frac{\binom{a}{x}}{\binom{N}{n}}$ we put the equivalent expression $\frac{n \cdot a}{N} \frac{\binom{a-1}{x-1}}{\binom{N-1}{n-1}}$ in all the terms above.

We further see that the expression on the right hand side is $\epsilon \cdot EX$, and in the hypergeometric distribution conserned $EX = \frac{a\,n}{N}$. Infering this as well, we get:

$$\frac{a\,n}{N} \sum_{x=0}^{k-1} \frac{\binom{a-1}{x-1}\binom{N-a}{n-x}}{\binom{N-1}{n-1}} + \frac{a\,n}{N} \cdot \gamma_1 \cdot \frac{\binom{a-1}{k-1}\binom{N-a}{n-x}}{\binom{N-1}{n-1}} + \frac{a\,n}{N} \cdot \gamma_2 \cdot \frac{\binom{a-1}{h-1}\binom{N-a}{n-x}}{\binom{N-1}{n-1}} +$$

$$+ \sum_{x=h+1}^{n} \frac{\binom{a-1}{x-1}\binom{N-a}{n-x}}{\binom{N-1}{n-1}} = \frac{a\,n}{N} \cdot \epsilon$$

Here $\frac{a\,n}{N}$ is a common factor, which then may be removed, since we assume $a > 0$, $n > 0$. Now, changing the summation variables from x to $y = x-1$ we get:

$$\sum_{y=o}^{k-2} \frac{\binom{a-1}{y}\binom{N-1-(a-1)}{n-1-y}}{\binom{N-1}{n-1}} + \gamma_1 \cdot \frac{\binom{a-1}{k-1}\binom{N-1-(a-1)}{n-1-(k-1)}}{\binom{N-1}{n-1}} + \gamma_2 \cdot \frac{\binom{a-1}{h-1}\binom{N-1-(a-1)}{n-1-(h-1)}}{\binom{N-1}{n-1}} +$$

$$+ \sum_{y=h}^{n-1} \frac{\binom{a-1}{y}\binom{N-1-(a-1)}{n-1-y}}{\binom{N-1}{n-1}} = \varepsilon$$

And substituting back to the original variables, using c instead of y, we get:

$$\begin{aligned}
&\sum_{c=o}^{k_i(T_i)-2} \frac{\binom{v_{r-i}-1}{c}\binom{v_{r-i-1}-1-(v_{r-i}-1)}{c_{r-i-1}-1-c}}{\binom{v_{r-i-1}-1}{c_{r-i-1}-1}} + \\
&+ \gamma_{1i}(T_i) \cdot \frac{\binom{v_{r-i}-1}{k_i(T_i)-1}\binom{v_{r-i-1}-1-(v_{r-i}-1)}{c_{r-i-1}-1-(k_i(T_i)-1)}}{\binom{v_{r-i-1}-1}{c_{r-i-1}-1}} + \\
&+ \gamma_{2i}(T_i) \cdot \frac{\binom{v_{r-i}-1}{h_i(T_i)-1}\binom{v_{r-i-1}-1-(v_{r-i}-1)}{c_{r-i-1}-1-(h_i(T_i)-1)}}{\binom{v_{r-i-1}-1}{c_{r-i-1}-1}} + \\
&+ \sum_{c=h_i(T_i)}^{c_{r-i-1}-1} \frac{\binom{v_{r-i}-1}{c}\binom{v_{r-i-1}-1-(v_{r-i}-1)}{c_{r-i-1}-1-c}}{\binom{v_{r-i-1}-1}{c_{r-i-1}-1}} = \varepsilon_i
\end{aligned} \qquad (2.31)$$

Then the system (2.21) is transformed in such a way that it may be solved by using tables of the hypergeometric distributions in terms of $k_i(T_i)$, $h_i(T_i)$, $\gamma_{1i}(T_i)$, $\gamma_{2i}(T_i)$.

2.4 The test-procedure following from the principle of retained parameter space

We have the same test-situation as in 2.3, but now the procedure is different. We will not, as in 2.3, reduce the parameter-space when a sub-hypothesis is accepted. Instead, we restrict the observationspace on the next step to the vectors leading to accepting the present sub-hypothesis, or equivalently, on the next step we will consider the conditional distribution of the observation-variables, given that the hypothesis on the present step is accepted. The hypothesis of independence in the $2 \times r$ table is rejected if (and only if) one of the sub-hypothesis is rejected, and hence accepted if all sub-hypothesis are accepted.

As for the method in 2.3, we would like to use randomized test-functions for the sub-hypothesis. This, however, leads to serious problems of calculation, because of the conditioning with respect to the acceptance of the former sub-hypothesis on each step.

Assume namely that the test found on the first step is randomized. Then we have exactly the same situation as in 2.3, and hence the test is given by (2.16) with the associated conditions on the constants, which are determined by means of the distribution (2.19a) found above.

On the second step we shall now consider the conditional distribution given that the first sub-hypothesis is accepted. The probatility of this event is easily found to be $1-\varepsilon$, under the null-hypothesis, and generally it may shortly be written $1-\beta_1(\varkappa)$ where $\beta_1(\varkappa)$ is the powerfunction of the test on the first step.

The conditional distribution in question is then found by dividing (2.14) by $1-\beta_1(\varkappa)$ to be

$$dP' = (2rp_1)^N \frac{1-\epsilon_1}{1-\beta_1(\varkappa)} e^{\sum_{j=1}^{r-1}\varkappa_i c_i + \tau_o W + \sum_{j=1}^{r-1}\tau_j v_j} dP'_o \; ; \quad T \in A_1$$

where dP'_o has the following form by the randomization on step 1 :

$$dP'_o = \begin{cases} \dfrac{1-\gamma_{11}(t_1)}{1-\epsilon_1} k_N(t)(\frac{1}{2r})^N d\mu; & \text{for } T: C_{r-1} = k_1(T_1) \\ \dfrac{1}{1-\epsilon_1} k_N(t)(\frac{1}{2r})^N d\mu; & \text{for } T: k_i(T_1) < C_{r-1} < h_1(T_1) \\ \dfrac{1-\gamma_{21}(t_1)}{1-\epsilon_1} k_N(t)(\frac{1}{2r})^N d\mu; & \text{for } T: C_{r-1} = h_1(T_1) \end{cases} \quad (2.40)$$

where now $k_N(t)$ is the coefficient $k_N(x,y)$ expressed in the new variables $T = (C_1,\dots,C_{r-1},W,V_1,\dots,V_{r-1})$.

The next step would be to find the test-procedure on step 2. Since dP' is D-K , the test is found in the usual way, analogous to (2.20), where T_2 is replaced by $T_2^* = (C_1,\dots,C_{r-3},C_{r-1},W,V_1,\dots,V_{r-1})$. To determine the constants, we would need the conditional distribution of C_{r-2} given T_2^* (from the P'_o-distribution). This is found by dividing dP'_o by the expression found by summing dP'_o over all possible C_{r-2}-values.

Here the condition $T: C_{r-1} = k_1(T_1)$ in (2.40) means all T such that the component C_{r-1} is equal to the lower rejection limit $k_1(T_1)$, which depends both on the level and on T .

Now assume that on step 1 we observe $C_{r-1} = k_1(T_1)$ and that the random-drawing leads to acceptance of the sub-hypothesis. Then we have

$$dP'_o = \frac{1-\gamma_{11}(t_1)}{1-\epsilon_1} k_N(t)(\frac{1}{2r})d\mu \; .$$

When we now want to sum over all C_{r-2}, we realize that C_{r-2} is a component of T_1, which means that $\gamma_{11}(T_1)$ varies with C_{r-2}. Then we must compute the $\gamma_{11}(T_1)$ for all these T_1 values, and this may take a considerable amount of time in practice.

The different forms of the dP_0'' (the conditional distribution on step 3) will be 9, that is 3 times the number on step 2. It is therefore very difficult, if possible, to find these randomized tests, eventhough it might be possible in practice for small values of N and a lot of available computertime.

We shall therefore restrict ourselves to non-randomized test-procedures for each sub-hypothesis. This means that we lose the exact level of the testprocedures, and hence the possibility to compare with the testprocedures of 2.3. But it will still be of interest to find the testprocedure.

Step 1.

Here the test-situation is again exactly equal to the one in 2.3, hence the test-procedure is given by (2.16) with $\gamma_{11} = \gamma_{21} = 0$, and the constants k_1 and h_1 are to be determined as the smallest and largest integer respectively such that:

$$\begin{aligned} &E_0(\delta_1^* \mid T_1) \leq \epsilon_1 \\ &E_0(\delta_1^* C_{r-1} \mid T_1) \leq \epsilon_1 \cdot E_0(C_{r-1} \mid T_1) \end{aligned} \tag{2.41}$$

where now δ_1^* is the test in (2.16) with $\gamma_{11} = \gamma_{21} = 0$ and where $E_0(\cdot \mid T_1)$ means the expectation in (2.19a), that is given T_1.

Note that the inequality in the first equation follows from the fact that the underlying distribution is discrete,/hence an equality cannot always be ecactly fulfilled. The inequality in the second line is chosen as a convention.

The level of this test is thus $\leq \epsilon$, as a matter of fact it is equal to $E_0(\delta_1^* | T_1)$, which is a function of T_1. Hence the level is a random variable.

Step 2.

We now consider the contitional distribution of T given that the sub-hypothesis on step 1 is accepted. In this distribution we will test $H_2 : \varkappa_{r-2} = 0$ as before. Now write $T_1^* = T_1$ to distinguish from the names in 2.3. Let the possible values of T_1^* be the vectors in

$$D_1 = \{t_1^{(1)}, t_1^{(2)}, \ldots, t_1^{(n_1)}\} \quad , \qquad \text{where} \quad n_1 < \infty .$$

and let the conditional acceptance-region for H_1 given $T_1 = t_1^{(j)}$ be

$$A_1(t_1^{(j)}) = \{(c_{r-1}, t_1^{(j)}) | k_1(t_1^{(j)}) \leq c_{r-1} \leq h_1(t_1^{(j)})\} \qquad (2.42)$$

where of course $k_1(T_1^*)$ and $h_1(T_1^*)$ are the rejection-numbers of the test (to be precise, we should have written $k_1^*(T_1^*)$) given by (2.16) by $\gamma_{11} = \gamma_{21} = 0$ and (2.41).

Hence the unconditional acceptance region is the union over D_1 of all $A_1(t_1^{(j)})$, that is

$$A_1 = \bigcup_{t_1^{(j)} \in D_1} A_1(t_1^{(j)}) \quad {}^{*)} \qquad (2.43)$$

when the hypothesis is correct, we find

$$Pr(T \in A_1) = Pr((C_{r-1}, T_1^*) \in A_1) = \sum_{t_1^{(j)} \in D} Pr(C_{r-1} \in A_1(t_1^{(j)}) | T_1^* = t_1^{(j)}) Pr(T_1^* = t_1^{(j)}) =$$

$$= \sum_{t_1^{(j)} \in D_1} (1 - \epsilon_1(t_1^{(j)})) \cdot Pr(T_1^* = t_1^{(j)}) ,$$

*) The A_i not to be mixed up with the A_i, the i^{th} level of the factor A ; $i = 1, 2, \ldots, r-1$.

where $\epsilon_1(t_1^{(j)})$ is the actual level of the test δ_1^* , when $T_1^* = t_1^{(j)}$.

The conditional distribution of $T|T \in A_1$, is now found by dividing (2.14) by $\Pr(T \in A_1)$, and may be put on D-K form:

$$dP^{A_1} = k_1(\varkappa,\tau) e^{\sum_{i=1}^{r-1} \varkappa_i c_i + \tau_0 W + \sum_{j=1}^{r-1} \tau_j v_j} dP_0^A \quad ; \text{ for } T \in A_1 \tag{2.44}$$

where $$dP_0^{A_1} = \frac{1}{\Pr(T \in A_1)} k_N(t) \left(\frac{1}{2r}\right)^N d\mu \quad ; \quad t \in A_1 \tag{2.45}$$

according to (2.15).

Since (2.44) is a D-K class, the test for H_2 is given by:

$$\delta_2^*(C_{r-2};T_2^*) = \begin{cases} 1 & \text{if } C_{r-2} < k_2(T_2^*) \text{ or } C_{r-2} > h_2(T_2^*) \\ 0 & \text{otherwise} \end{cases} \tag{2.46}$$

where $T_2^* = (C_1, C_2, \ldots, C_{r-3}, C_{r-1}, W, V_1, \ldots, V_{r-1})$, and where the constants k_2, h_2 (functions of T_2^*) are determined as the smallest and largest intergers respectively fulfilling:

$$\begin{aligned} &E_0(\delta_2^* | T_2^*) \leq \epsilon_2 \\ &E_0(\delta_2^* \cdot C_{r-2} | T_2^*) \leq \epsilon_2 \cdot E_0(C_{r-2} | T_2^*) \end{aligned} \tag{2.47}$$

where $E_0(\cdot | T_2^*)$ are the conditional expectation given T_2^* in $P_0^{A_1}$.

To determine the constants, we need $dP_0^{A_1}$. From (2.18) and (2.45), this distribution is given by:

$$\frac{1}{\Pr(T \in A_1)} \cdot \frac{N!}{(N-v_1-(w-c_1))!(v_1-v_2-(c_1-c_2))! \ldots (v_{r-1}-c_{r-1})!} \cdot$$
$$\cdot \frac{1}{(w-c_1)!(c_1-c_2)! \ldots c_{r-1}!} \cdot \left(\frac{1}{2r}\right)^N \tag{2.48}$$

Finally, to find the conditional distribution of C_{r-2} given T_2^*, we need the marginal distribution of T_2^*. This is as usual found by summing (2.48) over all C_{r-2}, where T is restricted to A_1. Is it then certain that the values of C_{r-2} restricted to $T \in A_1$ are consecutive integers? A priori we know that $C_{r-2} \in [C_{r-1}, C_{r-3}]$ from the definition of the C_i's. It is now obvious that the observed value of T is in A_1. Let us therefore keep all the components except C_{r-2} in this value fixed. Is it then possible for C_{r-2} to assume every integer value in the interval $[C_{r-1}, C_{r-3}]$ without bringing T out of A_1?

To answer this question, consider the definition of the C_j's as

$$
\begin{aligned}
C_{r-1} &= Y_r \\
C_{r-2} &= Y_r + Y_{r-1} \\
C_{r-3} &= Y_r + Y_{r-1} + Y_{r-2} \\
&\vdots
\end{aligned}
$$

Now all the C_i-values except C_{r-2} are fixed, therefore Y_r is fixed. Further $C_{r-3} = Y_r + Y_{r-1} + Y_{r-2}$ is fixed while Y_{r-1} may vary freely between 0 and $C_{r-3} - Y_r$ (in the first case $Y_{r-2} = C_{r-3} - Y_r$, in the second is $Y_{r-2} = 0$) without C_{r-1} and C_{r-3} (or the other C_i-values) being changed. But this means that C_{r-2} may assume every integer value between C_{r-1} and C_{r-3}.

Hence we may sum (2.48) over all values of C_{r-2}. Again, we first sum the terms containing C_{r-2}:

$$\sum_{c_{r-2}=c_{r-1}}^{c_{r-3}} \frac{1}{(v_{r-3}-v_{r-2}-(c_{r-3}-c_{r-2}))!(v_{r-2}-v_{r-1}-(c_{r-2}-c_{r-1}))!(c_{r-3}-c_{r-2})!(c_{r-2}-c_{r-1})!} =$$

$$= \sum_{c_{r-2}=c_{r-1}}^{c_{r-3}} \binom{v_{r-3}-v_{r-2}}{c_{r-3}-c_{r-2}}\binom{v_{r-2}-v_{r-1}}{c_{r-2}-c_{r-1}} \frac{1}{(v_{r-3}-v_{r-2})!(v_{r-2}-v_{r-1})!} =$$

$$= \binom{v_{r-3}-v_{r-1}}{c_{r-3}-c_{r-1}} \frac{1}{(v_{r-3}-v_{r-2})!(v_{r-2}-v_{r-1})!} \quad ,$$

using a well-known property of the hypergeometric distribution.

Now we find the desired distribution $g_2^*(c_{r-2}, t_2)$ by dividing (2.48) by the expression obtained by summing (2.48) over all c_{r-2}. We carry out the division, removing equal terms in numerator and denominator (i.e. those terms not containing c_{r-2}). Then the result becomes:

$$\frac{\dfrac{1}{(v_{r-3}-v_{r-2}-(c_{r-3}-c_{r-2}))!(v_{r-2}-v_{r-1}-(c_{r-2}-c_{r-1}))!(c_{r-3}-c_{r-2})!(c_{r-2}-c_{r-1})!}}{\dbinom{v_{r-3}-v_{r-1}}{c_{r-3}-c_{r-1}} \dfrac{1}{(v_{r-3}-v_{r-2})!(v_{r-2}-v_{r-1})!}} = \tag{2.49}$$

$$= g_2^*(c_{r-2};t_2) = \frac{\binom{v_{r-3}-v_{r-2}}{c_{r-3}-c_{r-2}}\binom{v_{r-2}-v_{r-1}}{c_{r-2}-c_{r-1}}}{\binom{v_{r-3}-v_{r-1}}{c_{r-3}-c_{r-1}}}, \quad \text{for } T \in A_1 \text{ and } c_{r-2} = c_{r-1}, c_{r-1}+1, \ldots, c_{r-3} .$$

Thus everything is known, and the test on step 2 may be found explicitely.

Step i.

We assume that step i - 1 is accepted, and we want to test

$H_i: \varkappa_{r-i} = 0$ against a two-sided alternative in the conditional distribution of T given that $T \in A_{i-1}$, the unconditional acceptance region on step i - 1. This region may be found analoguously to the region A_1 on step 2. We have

$$A_{i-1} = \bigcup_{t_{i-1}^{(j)} \in D_{i-1}} A_{i-1}(t_{i-1}^{(j)}) , \tag{2.50}$$

where $D_{i-1} = \left\{ t_{i-1}^{(1)}, t_{i-1}^{(2)}, \ldots, t_{i-1}^{(n_{i-1})} \right\}$ are the possible values of T restricted to $A_1 \cap A_2 \cap \ldots \cap A_{i-2}$, that is the hypothesis $H_1, \ldots, H_{i-1}$ are accepted. Here

$$A_{i-1}(t_{i-1}^{(j)}) = \left\{ (c_{r-i+1}, t_{i-1}^{(j)}) \mid k_{i-1}(t_{i-1}^{(j)}) \leq c_{r-i+1} \leq h_{i-1}(t_{i-1}^{(j)}) \right\} \tag{2.51}$$

is the conditional acceptance region on the (i-1)th step given that $T = t_{i-1}^{(j)}$ (and that $T \in A_1 \cap \ldots \cap A_{i-2}$).

Obviously $A_1 \cap \ldots \cap A_k = A_k$, since we on each each consider the observation space on the former step, restricted to the values of T that lead to acceptance of the present sub-hypothesis.

Again we find that

$$\begin{aligned} Pr(T \in A_{i-1}) &= Pr((C_{r-i+1}, T_{i-1}^{*}) \in A_{i-1}) = \\ &= \sum_{t_{i-1}^{(j)} \in D_{i-1}} Pr(C_{r-i+1} \in A_{i-1}(t_{i-1}^{(j)}) \mid T_{i-1}^{*} = t_{i-1}^{(j)}) \cdot Pr(T_{i-1}^{*} = t_{i-1}^{(j)}) \end{aligned} \tag{2.52}$$

where of course all the probabilities are conditioned by $T \in A_{i-2}$ (so that unconditional acceptance region here means unconditional relative to T_{i-1}^{*}. We must always keep, the condition $T \in A_{i-2}$ in mind).

The conditional distribution of T given $T \in A_{i-1}$ ($\cap A_{i-2} \cap \ldots \cap A_1$) now is found by division of the corresponding conditional distribution on the (i-1)th step by $Pr(T \in A_{i-1} \mid T \in A_{i-2})$, that is by (2.52). Thus the conditional distribution of $T|T \in A_{i-1}$ may be written on D-K form:

$$dP^{A_{i-1}} = K_{i-1}(\varkappa,\tau)e^{\sum_{i=1}^{r-1}\varkappa_i c_i+\tau_0 w+\sum_{j=1}^{r-1}\tau_j v_j}dP_0^{A_{i-1}}; \quad T \in A_{i-1} \tag{2.53}$$

where

$$dP_0^{A_{i-1}} = \frac{1}{Pr(T \in A_1)\ Pr(T \in A_2|T \in A_1)\dots\ Pr(T \in A_{i-2})}\cdot k_N(t)\left(\frac{1}{2r}\right)^N d\mu \tag{2.54}$$

This way of reasoning is of course inductive, but there is no reason to carry the whole procedure through to show (2.54), since this was done from step 1 to step 2. That (2.53) becomes a D-K class is evident, a D-K function is divided by a constant which is incorporated in $dP_0^{A_{i-1}}$, hence $K_i(\varkappa,\tau) = K_1(\varkappa,\tau)$ $i = 1,2,\dots,r-1$.

This leads to the test for H_i: $\varkappa_{r-i} = 0$ being found in the usual way, this time as a non-randomized test we find

$$\delta_i^*(C_{r-i};T_i^*) = \begin{cases} 1 \text{ when } & C_{r-i} < k_i(T_i^*) \text{ or } C_{r-i} > h_i(T_i^*) \\ 0 & \text{otherwise} \end{cases} \tag{2.55}$$

where now $T_i^* = (C_1,\dots,C_{r-i-1},C_{r-i+1},\dots,C_{r-1},W,V_1,\dots,V_{r-1})$ and the constants k_i, h_i (functions of T_i^*) is to be determined as the smallest and largest integer, respectively, such that

$$\begin{aligned} &E_0(\delta_i^*|T_i^*) \leq \epsilon_i \\ &E_0(\delta_i^* C_{r-i}|T_i^*) \leq \epsilon_i \cdot E_0(C_{r-i}|T_i^*) \end{aligned} \tag{2.56}$$

where $E_0(\cdot|T_i^*)$ is the conditional expectation giving T_i^* in distribution $P_0^{A_{i-1}}$.

Now let:

$$g_i^*(c_{r-i};t_i^{(j)}) = \Pr(C_{r-i} = c_{r-i} | T_i^* = t_i^{(j)}) =$$

$$= \frac{\Pr(C_{r-i} = c_{r-i} \cap T_i^* = t_i^{(j)})}{\Pr(T_i^* = t_i^{(j)})} = \frac{f^{(i-1)}(c_{r-i};t_i^{(j)})}{f_1^{(i-1)}(t_i^{(j)})} \tag{2.57}$$

where $D_i = \{t_i^{(1)},\ldots,t_i^{(n_i)}\}$ is the possible values of T given $T \in A_{i-1}$.

As under step 2, (2.18) and (2.54) leads to

$$f^{(i-1)}(c_{r-i},t_i^{(j)}) = \frac{1}{\Pr(T\in A_1)\Pr(T\in A_2|T\in A_1)\ldots \Pr(T\in A_{i-1}|T\in A_{i-2})} \cdot$$

$$\cdot \frac{N!}{(N-v_1-(w-c_1))!(v_1-v_2-(c_1-c_2))!\ldots (v_{r-1}-c_{r-1})!} \cdot$$

$$\cdot \frac{1}{(w-c_1)!(c_1-c_2)!\ldots c_{r-1}!} \left(\frac{1}{2r}\right)^N , \tag{2.58}$$

where $f^{(i-1)}(c_{r-i},t_i^{(j)})$ is defined in (2.57), and where $t_i^{(j)}$ defines $(C_1,\ldots,C_{r-i-1},C_{r-i+1},\ldots,C_{r-1},W,V_1,\ldots,V_{r-1})$.

Now $f_1^{(i-1)}(t_i^{(j)})$ is found by summing $f^{(i-1)}(c_{r-i},t_i^{(j)})$ over all values of C_{r-i}, and as under step 2, it can be shown that these are $C_{r-i+1},C_{r-i+1}+1,\ldots,C_{r-i-1}-1,C_{r-i-1}$. The procedure of summing is carried out the normal way, giving

$$f_1^{(i-1)}(t_i^{(j)}) = \frac{1}{\Pr(T\in A_1)\cdot\Pr(T\in A_2|T\in A_1),\ldots.\Pr(T\in A_{i-1}|T\in A_{i-2})} \cdot$$

$$\cdot \frac{N!}{(N-v_1-(w-c_1))!(v_1-v_2-(c_1-c_2))!\ldots (v_{r-i-2}-v_{r-i-1}-(c_{r-i-2}-c_{r-i-1}))!} \cdot$$

$$\cdot \frac{1}{(v_{r-i+1}-v_{r-i+2}-(c_{r-i+1}-c_{r-i+2}))!\ldots (v_{r-1}-c_{r-1})!} \cdot \tag{2.58}$$

$$\cdot \frac{1}{(w-c_1)!(c_1-c_2)!\ldots(c_{r-i-2}-c_{r-i-1})!(c_{r-i+1}-c_{r-i+2})!\ldots c_{r-1}!}\left(\frac{1}{2r}\right)^N .$$

$$\cdot \frac{1}{(v_{r-i-1}-v_{r-i})!(v_{r-i}-v_{r-i+1})!}\binom{v_{r-i-1}-v_{r-i+1}}{c_{r-i-1}-c_{r-i+1}} \tag{2.58}$$

and finally, dividing (2.57) by (2.58), cancelling equal terms, we get

$$\frac{\dfrac{1}{(v_{r-i-1}-v_{r-i}-(c_{r-i-1}-c_{r-i}))!(v_{r-i}-v_{r-i+1}-(c_{r-i}-c_{r-i+1}))!(c_{r-i-1}-c_{r-i})!(c_{r-i}-c_{r-i+1})!}}{\dfrac{1}{(v_{r-i-1}-v_{r-i})!(v_{r-i}-v_{r-i+1})!}\dbinom{v_{r-i-1}-v_{r-i+1}}{c_{r-i-1}-c_{r-i+1}}}$$

Rearranging terms and inferring binomial coefficients gives:

$$g_i^*(c_{r-i}, t_i^{(j)}) = \frac{\dbinom{v_{r-i-1}-v_{r-i}}{c_{r-i-1}-c_{r-i}}\dbinom{v_{r-i}-v_{r-i+1}}{c_{r-i}-c_{r-i+1}}}{\dbinom{v_{r-i-1}-v_{r-i+1}}{c_{r-i-1}-c_{r-i+1}}} \tag{2.59}$$

for $T \in A_{i-1}$; $C_{r-i} = C_{r-i+1}, C_{r-i+1}+1, \ldots, C_{r-i-1}$.

Then the test-procedure is fixed. For $i = r-1$, we just have to remember that $V_o = N$, $C_o = W$. Going back to the original variables (x_i, y_i) ; $i = 1,2,\ldots,r$ in (2.59), we get

$$\frac{\dbinom{u_{r-i}}{y_{r-i}}\dbinom{u_{r-i+1}}{y_{r-i+1}}}{\dbinom{u_{r-i}+u_{r-i+1}}{y_{r-i}+y_{r-i+1}}} ,$$

which gives us the following intuitive interpretation of the stepwise method:

On the first step we test for independence in the 2×2 table $(B,B^*) \times (A_{r-1},A_r)$; on the second in the 2×2 table $(B,B^*) \times$

$\times (A_{r-2}-A_{r-1}),\ldots$ on the $i^{\underline{th}}$ step in the 2×2 table $(B,B^*)\times$ $\times (A_{r-i},A_{r-i+1})$, and finally on the $(r-1)$th step we test for independence in the 2×2 table $(B,B^*)\times(A_1,A_2)$, where $A_1,\ldots,A_r$ are the different levels of A.

For every sub-hypothesis not rejected, we restrict the observation space on the next step to the values leading to acceptance of the hypothesis on that step.

Finally we make some considerations of the level of the test. Let E_i and F_i have the same meaning as in 2.3. Then we have

$$\epsilon = \Pr(\text{rejecting the hypothesis H}) = 1-\Pr(\text{accepting H}) =$$

$$= 1-\Pr(\bigcap_{i=1}^{r-1} \text{accepting}\,H_i) = 1-\Pr(\bigcap_{i=1}^{r-1} E_i) \quad \text{when } H \text{ is true.}$$

Then

$$\Pr(\bigcap_{i=1}^{r-1} E_i) = \Pr(E_1)\cdot\Pr(E_2|E_1)\ \ldots\ \Pr(E_{r-1}|E_1\cap\ldots\cap E_{r-2}) =$$

$$= \Pr(T\in A_1)\cdot\Pr(T\in A_2|T\in A_1)\ \ldots\ \Pr(T\in A_{r-1}|T\in A_1\cap\ldots\cap A_{r-2})\ .$$

since A_i was the acceptance region on the $i^{\underline{th}}$ step. Now the conditional level on the $i^{\underline{th}}$ step is less than or equal to ϵ_i for all values of T_i^*, hence the unconditional level of the test δ_i^* is $\leq \epsilon_i$, for all i. But then for the real total level ϵ', we have

$$\epsilon' \leq 1-(1-\epsilon_1)(1-\epsilon_2\cdot\ldots\cdot(1-\epsilon_{r-1})$$

Then reasoning is the same as in 2.3, hence putting $\epsilon_i = \alpha$ $i = 1,2,\ldots,r-1$, we find that

$$\epsilon_i = 1-\sqrt[r-1]{1-\epsilon} \quad ; \quad i = 1,2,\ldots,r-1$$

and $\epsilon' \leq \epsilon$ for all possible values of T. But the exactness of the inequality may be very bad in some situations.

We conclude this chapter by saying that the test seems to have little practical importance as long as the actual level is unknown. Practical applications of the method shows that when the randomization is removed, the level decreases considerably, especially when N is small. To obtain "good" test-methods on each step, we have used the principle of unbiasedness. If we are willing to weaken this condition, we may obtain tests with higher total (conditional) level. But these tests may be very skew, that is, approximately one-sided and hence tend to discover departures from the hypothesis in only one direction.

If one could find the exact level, that is, being able to calculate the randomized test-procedures on each step, the test might have practical applications.

3. Practical applications.

Both test-procedures which were developed in ch. 2.3 and 2.4 have been programmed for a computer, and some examples have been run. The programs were run at the computer installation of the University of Oslo, a CDC 3300.

We shall concentrate on the method of ch. 2.3. The applications of the other procedure is similar.

The observations are punched on cards together with some parameters telling the computer which method to apply, the level of the test and the dimension of the table, r . The main practical problem is to solve the two equations (2.29) and (2.31) simultanuously. To do this numerically, we have to rewrite the equations:

Now let

$$f_N(x;n,a) = \frac{\binom{a}{x}\binom{N-a}{n-x}}{\binom{N}{n}} \quad ; \; x = 0,1,\dots,\min[n,a]$$

and let

$$F_N(c;n,a) = \sum_{x=0}^{c} f_N(x;n,a) \; .$$

Then the equations (2.29) and (2.31) will correspond to equations of the type

$$F_N(c_1-1;n,a) + \gamma_1 \cdot f_N(c_1;n,a) + \gamma_2 \cdot f_N(c_2;n,a) + 1 - F_N(c_2;n,a) = \varepsilon \tag{3.1}$$

$$F_{N-1}(c_1-2;n-1,a-1) + \gamma_1 \cdot f_{N-1}(c_1-1;n-1,a-1) + \gamma_2 \cdot f_{N-1}(c_2-1;n-1,a-1) + 1 - F_{N-1}(c_2-1;n-1,a-1) = \varepsilon$$

where we have made the notations a little bit simpler.

The problem is to find integers $0 < c_1 < c_2 < n$ and real numbers γ_1, γ_2 $0 \leq \gamma_i < 1$; $i = 1,2$ which fulfill (3.1). The extra conditions are put on to give us a unique solution. So we have two equations with four unknowns, plus the side conditions.

These must be incorporated in the equations.

Now let $[\cdot]$ be the integer function, that is $[x]$ = greatest integer $\leq x$, and define G_N and H_N as follows:

$$\begin{aligned} G_N(x;n,a) &= F_N([x]-1;n,a) + (x-[x])f_N([x];n,a) \\ H_N(x;n,a) &= 1 - F_N([x-\delta]+1;n,a) + ([x-\delta]+1-x)f_N([x-\delta]+1;n,a) \end{aligned} \tag{3.2}$$

where $\delta = 0^+$, that is by definition: Let $x = n+\varkappa$; $0 \leq \varkappa < 1$. Then

$$[x-0^+] = \begin{cases} n-1 & \text{if } \varkappa = 0 \\ n & \text{if } 0 < \varkappa < 1 \end{cases} \tag{3.3}$$

Further let

$$\begin{aligned} x &= c_1 + \gamma_1 \\ y &= c_2 - \gamma_2 \end{aligned} \tag{3.4}$$

Then we can write:

$$\begin{aligned} c_1 &= [x] , & \gamma_1 &= x-c_1 = x-[x] \\ c_2 &= [y-0^+]+1 , & \gamma_2 &= [y-0^+]+1-y \end{aligned} \tag{3.5}$$

The second line in (3.5) is found as follows: We have $y = c_2-\gamma_2$, $0 \leq \gamma_2 < 1$. If we put $c_2 = [y+1]$, and if y is an integer, e.g. $y = m$, we get $\gamma_2 = [y+1]-y = m+1-m = 1$, which is an illegal value, while $c_2 = [y-0^+]+1$ gives $\gamma_2 = [m-0^+]+1-m = m-1+1-m = 0$, which is a legal value.

Using the results (3.5) and (3.1) we get:

$$\begin{aligned} &F_N([x]-1;n,a) + (x-[x])f_N([x];n,a) + \\ &\quad + ([y-0^+]+1-y)f_N([y-0^+]+1;n,a) + 1 - F_N([y-0^+]+1;n,a) = \varepsilon \\ &F_{N-1}([x]-2;n-1,a-1) + (x-[x])f_{N-1}([x]-1;n-1,a-1) + \\ &\quad + ([y-0^+]+1-y)f_{N-1}([y-0^+];n-1,a-1) + 1-F_{N-1}([y-0^+];n-1,a-1) = \varepsilon \end{aligned} \tag{3.6}$$

In (3.6) we infer G_N and H_N given by (3.2). This gives:

$$\begin{aligned} G_N(x;n,a) + H_N(y;n,a) &= \epsilon \\ G_{N-1}(x-1;n-1,a-1) + H_{N-1}(y-1;n-1,a-1) &= \epsilon \end{aligned} \tag{3.7}$$

which is a system of two equations with two unknowns x and y. From (3.2) we see that G_N and H_N are continuous functions of x. For write $G_N(x) = G_N(x;n,r)$. This function is continuous on all open intervals of the type $(m,m+1)$, where $m = 0,1,2,\ldots$ $\ldots,\min(n,a)+1$. We further have

$$G_N(m) = F_N(m-1;n,a) ,$$

and

$$\lim_{x \to m-} G_N(x) = F_N(m-2;n,a) + f_N(m-1;n,a) = F_N(m-1;n,a)$$

and finally

$$\lim_{x \to m+} G_N(x) = F_N(m-1;n,a)$$

which shows the continuity of G_N in the interval $[0,\min(a,n)+1]$. A similar argument can be given for H_N.

Then the system (3.7) may be solved numerically by a double iteration method. In the program, the Newton method with variable chord is used on the system

(i) $G_N(x;n,a) + H_N(y;n,a) - \epsilon = 0$

(ii) $G_{N-1}(x-1;n-1,a-1) + H_{N-1}(y-1;n-1,a-1) - \epsilon = 0$.

We first compute an x-value for a fixed y, such that (x,y) is a solution of (i). The x-value found is used in (ii) to compute a y-value, the new y-value is again used in (i) to find a new x-value, etc. The process may be shown to converge. During the calcultion, δ in (3.2) is given a very small value 10^{-9}. The same level is used on each step, that is $\epsilon_i = \alpha$

given by (2.28).

My intention was to compute some points of the power-function of the test-procedure in 2.3. To do this, however, the program should have to run through all a priori possible values of T_i. This number is increasing very rapidly, thus we have for a 2×2 table $a_N = \frac{(N+1)^2(N+2)}{2}$ possible values of T on the second step. On the first step, there are even more.

To show the point, assume $N = 100$, then we find $a_{100} = 520,251$. The calculation of a two-step test takes about 5 seconds, because of the iteration. (This calculation may perhaps be considerably reduced by using a more elaborate method). Multiplying by a_{100} gives about 2.5 million seconds, or about one month of continuous computer-time. Therefore the power-calculation was out of the question.

Therefore the examples lose a great deal of interest. What finally was done, was to compute the two tests described in ch.2 on some numerical examples - and to calculate the result of the classical test-procedure (the independence-test, see [6],pp.244-249). The examples show that our methods will discover some dependencies which are not found by the classical method - and vice verca. Especially, for $r \geq 4$, our methods may lead to rejection on level 0.05, while the critical value of the classical test-procedure is 0.20. This is true when the dependencies are concentrated on one place of the table.

We shall be very careful with drawing conclusions of the properties of the method, but it may seem that our methods are suitable for finding partial dependencies, which is often overlooked by the classical method since it tends to average over all the blocks of the table. On the other hand, the classical method may be better to find small, systematic dependencies.

As we have mentioned before, the method of the retained parameter space will get its real level considerable reduced, since we were unable to randomize. This mean we have very little background to tell the properties of the method. What one could try to do, is to compute the unconditional, actual level of the test, and compare with the other methods, using the level mentioned above. But this problem remains unsolved for two reasons. First, I have not been able to find a theoretic expression for the unconditional total level in terms at the conditional levels, because of the definition of T_i^*, which has the same dimension for every i, but different components.

Secondly, even if we had found an expression, we would meet the same problem as we mentioned for the power-calculations. The capacity of the computer.

Concerning the method of reduced parameter-space, an expression for the unconditional level is found relatively easy. Here we have T_i "is a part of " T_j for $i > j$, that is T_j consist of all the components of T_i in addition to some others. We shall find the unconditional level for $r = 3$.

On the first step, the conditional level given T_1 is called $\epsilon_1(T_1)$, and on the second the conditional level given T_2 is $\epsilon_2(T_2)$. As we mentioned above, T_2 is included in T_1, because

$$T_1 = (C_1, W, V_1, V_2) \ , \qquad T_2 = (W, V_1, V_2) \ ,$$

which means, that if T_1 is given, so is T_2. Then given T_1 is the same as given T_1 and T_2, that is $T_1 \cap T_2$. On the other hand if T_1 and T_2 are given, so is T_1. Hence

$$\epsilon_1(T_1) = \Pr(F_1|T_1) = \Pr(F_1|T_1 \cap T_2) = \Pr((F_1|T_1)|T_2) \qquad (3.8)$$

Then the conditional level on the first step, given T_2 is

$$\varepsilon_1'(T_2) = \sum_{t_1|t_2} \varepsilon_1(t_1) \cdot \Pr(T_1 = t_1 | t_2) \tag{3.9}$$

and hence the conditional total level given T_2 is by an analogue to (2.27)

$$\varepsilon(T_2) = \varepsilon_1'(T_2) + (1-\varepsilon_1'(T_2)) \cdot \varepsilon_2(T_2) \tag{3.10}$$

and finally we find the total unconditional level to be

$$\varepsilon = \sum_{t_2} \varepsilon(t_2) \cdot \Pr(T_2 = t_2) \tag{3.11}$$

Formulaes for $r > 3$ may be found in a similar way, but with considerably more writing.

These calculations are in fact unnecessary for finding the total (unconditional) level, since all the conditional levels are exactly ε_i on step i. But for power-calculation, the same formulaes are valid, we just replace the $\varepsilon_i(T_i)$ by $\beta_i(T_i)$, the conditional power given T_i. But, as we have seen, the power is too time-consuming to calculate.

To compute the method of retained parameter-space, we use the same procedure as for the method of reduced parameter-space, to solve the equation-system (3.7). We use the solution with $\gamma_1 = \gamma_2 = 0$ that is, the corresponding non-randomized test.

In the appendix, we show some examples which have been run on the computer. They should be self-explanatory.

Appendix

Here we enclose 5 examples run on a computer, using both methods. They are presented in the way the computer prints them.

For each step, one line is written, telling the value of C, the stochastic variable on which the test is based, and the rejection numbers K and H corresponding to $k_i(T_i)$ and $h_i(T_i)$ in the method of reduced parameter-space, and $k_i^*(T_i)$ and $h_i^*(T_i)$ in the method of retained parameter-space.

Finally, we calculate the result of the classical test-procedure, and also the critical level (that is, the least level on which rejection will be the result).

Problem 1.

Parameters: $r = 3$, $N = 52$.

Data:	X-values:	8	7	8
	Y-values:	4	10	15

Reduced parameter-space

Step 1: C = 15, K = 11, H = 18, No rejection, level: 0,0253

Step 2: C = 25, K = 19, H = 26, No rejection, level: 0,0253

Result: No rejection on level 0.05

Retained parameter-space

Step 1: C = 15, K = 11, H = 18, No rejection,level: 0.0072

Step 2: C = 25, K = 20, H = 26, No rejection, level: 0.0078

Result: No rejection, max level 0.05

The classical test-procedure gives no rejection on level 0.0500

(critical level 0.1877)

Problem 2.

Parameters: r = 5 , N = 380

Data:	x-values:	36	39	59	32	20
	y-values:	18	54	62	34	26

Reduced parameter-space

Step 1: C = 26, K = 18, H = 31, no rejection, level: 0.0127

Step 2: C = 60, K = 49, H = 68, no rejection, level: 0.0127

Step 3: C = 122,K = 116,H = 136,no rejection, level: 0.0127

Step 4: C = 176,K = 158,H = 175, rejection, level: 0.0127

Result: Rejection on step 4, level: 0.05.

Retained parameter-space

Step 1: C = 26, K = 18, H = 31, no rejection, level: 0.0068

Step 2: C = 60, K = 52, H = 68, no rejection, level: 0.0091

Step 3: C = 122, K = 117, H = 135, no rejection, level: 0.0088

Step 4: C = 176, K = 160, H = 175, rejection, level: 0.0059

Result: Rejection on step 4, max level 0.05

The classical test-procedure gives no rejection on level 0.05

(critical level 0.0569)

Problem 3.

Parameters: r = 5, N = 502

Data:	x-values:	50	35	22	33	41
	y-values:	69	98	49	47	58

Reduced parameter-space

Step 1: C = 58, K = 50, H = 66, no rejection, level: 0.0127

Step 2: C = 105, K = 102, H = 119, no rejection, level: 0.0127

Step 3: C = 154, K = 154, H = 176, no rejection, level: 0.0127

Step 4: C = 252, K = 234, H = 257, no rejection, level: 0.0127

Result: No rejection, level: 0.05

Retained parameter-space

Step 1: C = 58, K = 50, H = 66, no rejection, level: 0.0093

Step 2: C = 105, K = 102, H = 116, no rejection, level: 0.0108

Step 3: C = 154, K = 149, H = 164, no rejection, level: 0.0090

Step 4: C = 252, K = 233, H = 252, no rejection, level: 0.0072

Result: No rejection, max level 0.05

The classical test-procedure gives rejection on level 0.05

(critical level 0.0362)

Problem 4.

Parameters: r = 3, N = 100

Data:	x-values:	3	15	22
	y-values:	17	18	25

Reduced parameter-space

Step 1: C = 25, K = 20, H = 30, no rejection, level: 0.0253

Step 2: C = 43, K = 44, H = 53, rejection, level: 0.0253

Result: Rejection on step 2, level 0.05

Retained parameter-space

Step 1: C = 25, K = 20, H = 30, no rejection, level: 0.0123

Step 2: C = 43, K = 43, H = 51, no rejection, level: 0.0063

Result: No rejection, level 0.05

The classical test-procedure gives rejection on level 0.05

(critical level 0.0383)

Problem 5.

Parameters: $r = 4$, $N = 100$

Data:	x-values:	7	12	16	15
	y-values:	19	10	11	10

Reduced parameter-space

Step 1: C = 10, K = 6 , H = 14, no rejection, level: 0.0170

Step 2: C = 21, K = 17, H = 26, no rejection, level: 0.0170

Step 3: C = 31, K = 32, H = 32, rejection, level: 0.0170

Result: Rejection on step 3, level 0.0500

Retained parameter-space

Step 1: C = 10, K = 6 , H = 14, no rejection, level: 0.0102

Step 2: C = 21, K = 17, H = 26, no rejection, level: 0.0032

Step 3: C = 31, K = 30, H = 38, no rejection, level: 0.0077

Result: No rejection, level 0.05

Classical test-procedure gives no rejection at level 0.05

(critical level 0.0539)

References

[1] Erling Sverdrup: "Multiple decision theory". Lecture Notes Series No 15, Aarhus Universitet, 1969.

[2] Erling Sverdrup: "Statistical Inference problems for Darmois-Koopman Classes of Distributions". Matematisk institutt, Oslo Universitet, 1969.

[3] T.W. Anderson: "The choice of degree of a polynomial regression as a multiple decision problem". Am. Math. Stat., vol. 33 (1962) p. 255.

[4] E.L. Lehmann: "A theory of multiple decision problems II". Am. Math. Stat., vol. 28 (1957) p. 547.

[5] C.E. Frøberg: "Lärobok i numerisk analys". Svenska bokförlaget/Bonniers, Stockholm 1962, chap. 2.2 and 2.7.

[6] Erling Sverdrup: "Laws and chance variations", vol II, North-Holland Publishing Company, Amsterdam 1967.

[7] E.L. Lehmann: "Testing statistical hypothesis". John Wiley & Sons Inc., New York 1966.